

AMENDMENT TO THE CLAIMS

1. (Currently Amended) A speech processing system, comprising:
 - an acoustic model;
 - a model authoring component receiving an input schema that describes semantic classes and slots that model input words in a domain of interest.
 - a composite language model that supports a vocabulary of words and including a rules-based model portion that has a plurality of automatically generated grammar rules, the grammar rules being automatically generated by the model authoring component from ~~an~~ the input schema to define a rules-based grammar parse tree that maps words in a natural language speech input into portions of the rules-based grammar parse tree, and a statistical model portion having a plurality of statistical n-gram models trained based on training data, one statistical n-gram model corresponding to each of a plurality of pre-terminals, and wherein words in the vocabulary that are not used to train a specific statistical n-gram model comprise unseen words for the specific statistical n-gram model, the statistical model portion further comprising a backoff model n-gram, separate from the plurality of statistical n-gram models corresponding to the pre-terminals, which, when accessed, is configured to assign a backoff score to a word in the vocabulary, wherein each statistical n-gram model includes a reference to the backoff model portion for all unseen words; and
 - a decoder coupled to the acoustic model and the composite language model and configured to map portions of the natural language speech input to the pre-terminals and slots, derived from a schema, based on the acoustic model and the composite language model.
2. (Original) The speech processing system of claim 1 wherein the decoder is configured to map portions of the natural language speech input to the slots based on the rules-based model portion of the composite language model.

3. (Original) The speech processing system of claim 1 wherein the decoder is configured to map portions of the natural language speech input to the pre-terminals based on the statistical model portion of the composite language model.

4. Canceled.

5. Canceled.

6. Canceled.

7. Canceled.

8. (Previously Amended) The speech processing system of claim 1 wherein the backoff model n-gram assigns a uniform score to every word in the vocabulary.

9. (Original) The speech processing system of claim 1 wherein the rules-based model portion comprises:

a context free grammar (CFG).

10. (Currently Amended) A method of assigning probabilities to word hypotheses during speech processing performed by a computer with memory, comprising:

receiving a word hypothesis;

accessing a composite language model having a plurality of statistical models and a plurality of rules-based models;

assigning, with the processor, an n-gram probability, with an n-gram model, to the word hypothesis if the word hypothesis corresponds to a word seen during training of the n-gram model; and

referring to a separate backoff model for the word hypothesis if the word hypothesis corresponds to a word unseen during training of the n-gram model; and

assigning, with the processor, a uniform backoff probability to ~~each~~ every word

hypothesis, that corresponds to an unseen word, with the backoff model.

11. (Original) The method of claim 10 and further comprising:
mapping the word hypotheses to slots derived from an input schema based on the rules-based models in the composite language model.
12. (Original) The method of claim 11 and further comprising:
mapping the word hypotheses to pre-terminals derived from the input schema based on probabilities assigned by the n-gram models and the backoff model in the composite language model.
13. (Original) The method of claim 12 wherein referring to a separate backoff model comprises:
referring to a uniform distribution n-gram.
14. (Original) The method of claim 13 wherein assigning a backoff probability comprises:
assigning a uniform distribution score to every word in the vocabulary.
15. (Currently Amended) A composite language model for use in a speech recognition system, comprising:
an authoring component receiving a schema describing semantic classes and slots that model expected word inputs in a domain of interest, the authoring component generating an automatically learned rules-based model portion having automatically learned grammar rules, automatically generated by the authoring component from a ~~the~~ schema, to define a grammar that is accessed to map words in an input speech signal to portions of a rules-based grammar parse tree that has slots derived from the schema; ~~and~~ a statistical model portion accessed to map portions of the input speech signal to pre-terminals in the rules-based grammar parse tree derived from the schema; and

a computer processor, being a functional element of the composite language model
authoring system, activated by the authoring component to facilitate automatic
generation of the grammar rules

16. (Currently Amended) The composite language model authoring system of claim 15 wherein the statistical model portion comprises:

a plurality of statistical n-gram models, one statistical n-gram model corresponding to each pre-terminal.

17. (Currently Amended) The composite language model authoring system of claim 15 wherein the rules-based model portion comprises:

an automatically learned context free grammar (CFG), learned from an example base of training data examples.

18. (Currently Amended) The composite language model authoring system of claim 16 wherein the composite language model supports a vocabulary of words and wherein the statistical n-gram models are trained based on training data, and wherein words in the vocabulary that are not used to train a specific statistical n-gram model comprise unseen words for the specific statistical n-gram model.

19. (Currently Amended) The composite language model authoring system of claim 18 wherein the statistical model portion of the composite language model further comprises:

a backoff model portion which, when accessed, is configured to assign a backoff score to a word in the vocabulary.

20. (Currently Amended) The composite language model authoring system of claim 19 wherein each statistical n-gram model includes a reference to the backoff model portion for all unseen words.

21. (Currently Amended) The composite language model authoring system of claim 20 wherein the backoff model portion comprises:

- a uniform distribution n-gram that assigns a uniform score to every word in the vocabulary.